

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
26 July 2001 (26.07.2001)

PCT

(10) International Publication Number  
**WO 01/53475 A2**

(51) International Patent Classification<sup>7</sup>: **C12N 15/00**

(21) International Application Number: PCT/IT01/00008

(22) International Filing Date: 12 January 2001 (12.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
RM2000A000021 17 January 2000 (17.01.2000) IT

(71) Applicant (*for all designated States except US*): **UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"** [IT/IT]; Piazzale Aldo Moro, 5, I-00185 Roma (IT).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **COGONI, Carlo** [IT/IT]; Policlinico Umberto I, Università Degli Studi Di Roma "La Sapienza", Dipartimento Biotechnologie Cellulari ed Ematologia, Viale Regina Elena, 324, I-00161 Roma (IT). **MACINO, Giuseppe** [IT/IT]; Policlinico Umberto I, Università Degli Studi Di Roma "La Sapienza", Dipartimento Biotechnologie Cellulari ed Ematologia, Viale Regina Elena, 324, I-00161 Roma (IT). **CATALANOTTO, Caterina** [IT/IT]; Policlinico Umberto I, Università Degli Studi Di Roma "La Sapienza", Dipartimento Biotechnologie Cellulari ed Ematologia, Viale Regina Elena, 324, I-00161

Roma (IT). **AZZALIN, Gianluca** [IT/IT]; Policlinico Umberto I, Università Degli Studi Di Roma "La Sapienza", Dipartimento Biotechnologie Cellulari ed Ematologia, Viale Regina Elena, 324, I-00161 Roma (IT).

(74) Agents: **CAPASSO, Olga** et al.; Ing. Barzanò & Zanardo Roma S.p.A., Via Piemonte, 26, I-00187 Roma (IT).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: ISOLATION AND CHARACTERIZATION OF A *N. CRASSA* SILENCING GENE AND USES THEREOF

(57) Abstract: An isolated nucleic acid molecule encoding for a protein characterized in that it has a silencing activity and comprises a domain responsible for dsRNA interference is disclosed; furthermore expression vectors suitable for the expression of said sequence in bacteria, plants, animals and fungi are disclosed; the invention refers also to organisms transformed by such vectors.



**WO 01/53475 A2**

ISOLATION AND CHARACTERIZATION OF A *N. CRASSA* SILENCING GENE  
AND USES THEREOF

5           The present invention relates to the isolation and characterization of a *Neurospora crassa* gene encoding for an essential activity in the co-suppression process and to uses and applications thereof in vegetal, animal and fungine fields.

10           The production of transgenic organisms is of large utility both in basic and applied biological research. The transgenic DNA is usually integrated in the genome and transferred as a Mendelian character. However, in various instances, the transgene introduction induces  
15   gene silencing phenomena (Flavell, R.B. 1994), i.e. the repression of the expression of the transgene itself and/or of one or more endogenous homologous genes.

          The gene silencing (suppression of gene expression) can act at two levels: transcriptional (trans-  
20   inactivation) where transgenes contain sequences homologous to the silenced gene promoter (Vaucheret, 1993); and post-transcriptional (co-suppression) which requires homologies between coding regions (Flavell, 1994; Stam et al., 1997; Baulcombe, 1996).

25           Generally the silencing induced by a transgene requires an almost complete sequence homology (from 70% to 100%) between transgene and silenced gene sequences (Elkind, 1990).

          In the *Neurospora crassa* filamentous fungus, during  
30   the vegetative phase, the presence of transgenes induces a post-transcriptional gene silencing phenomenon, named "quelling" (Cogoni et al., 1996).

By using the *al-1* gene (albino 1) (Schmidhauser et al., 1990) as silencing visual marker, many features of the phenomenon have been discovered (Cogoni et al., 1996). Particularly the *al-1* gene "quelling" in *Neurospora* is characterized in that: 1) the gene silencing is reversible further to the loss of transgene copies; 2) the reduction of mRNA basal level results from a post-transcriptional effect; 3) transgenes containing at least a region of 132 base pairs which is identical to the region encoding for the target gene are sufficient to induce the "quelling"; 4) the duplication of promoter sequences is ineffective to induce the silencing; 5) the "quelling" exhibits a dominant behavior in heterocarions containing both transgenic and untransformed nuclei, indicating the involvement of a trans-acting diffusible molecule among the nuclei; 6) the expression of an aberrant RNA transcribed by the transgenic locus is strictly correlated to silencing, suggesting that the "quelling" can be induced and/or mediated by a transgenic RNA molecule.

Therefore homologies between *Neurospora* silencing and plant co-suppression can be pointed out. The gene silencing in *Neurospora* is reversible, as result of transgenic copies instability during mitotic phase; in plants also the co-suppression reversion is associated with the reduction of transgene copy number, resulting from intra-chromosomal recombination during mitosis or meiosis (Mittelstein Scheid et al., 1994; Stam et al., 1997). Thus both in plants and in *Neurospora* the transgene presence is required to maintain the silencing. As in *Neurospora*, a decrease of the mRNA basal level of the silenced gene results from a post-transcriptional

mechanism (Dehio and Schell 1994; van Blokand et al., 1994; de Carvalho et al., 1995). Furthermore to induce the "quelling", transgenes must contain a portion of the silencing target gene coding sequence, being the promoter region ineffective. In plants coding regions with no promoter sequences can induce silencing (van Blokand et al., 1994) and, as in the "quelling", promoters or functionally active gene products are not required for the co-suppression.

One of the similarities between "quelling" and co-suppression in plants is that both mechanisms are mediated by diffusion factors. In *Neurospora* heterokaryotic strains, nuclei wherein the *albino-1* gene is silenced are able to induce the *al-1* gene silencing of the other not transformed nuclei, all sharing the same cytoplasmic environment (Cogoni et al., 1996). In plants the presence of a diffusion factor results from the fact that the co-suppression is effective in inhibiting the replication of Tobacco Etch Virus (TEV), a RNA virus with an exclusively cytoplasmic cycle. The occurrence of highly diffusible factors, which are effective to mediate the co-suppression, has been demonstrated using the grafting technique in tobacco (Palaqui et al., 1997), showing that silenced tobacco plants are able to transfer the silencing to non-silenced plants through grafting.

The fact that "quelling" and co-suppression share all these features suggests that mechanisms involved in post-transcriptional gene silencing in plants and in fungi can be evolved by an ancestral common mechanism.

Recently gene inactivation phenomena resulting from transgene introduction have been disclosed in animals. In *Drosophila melanogaster* the location of a transgene close

to heterochromatic centers results in a variegate expression (Wallrath and Elgin, 1995; Pirrotta, V., 1997). Similar expression profiles have been observed when the reference transgene is within tandem arrayed transposons, indicating that tandem repeats are effective to induce the chromatin condensation. (Dorer and Henikoff, 1994). Again in *Drosophila* Pal-Bhadra et al. (1997) have observed that the transgene introduction can lead to gene inactivation phenomena, similar to the co-suppression.

Gene silencing phenomena resulting from transegene sequence repeats have been disclosed recently in mammals.

Garrick et al. (1998) produced mouse transgenic lines wherein 100 transgenic copies are present in a unique locus and are repeats-arrayed in direct tandem. The transgene expression has been disclosed to be inversely proportional to the number of occurring copies, indicating that silencing phenomena dependent on repeat copies are present also in mammals.

It has been recently found that double stranded RNA molecules can induce a sequence-specific silencing in several organisms (Fire A., 1999). The mechanism known as dsRNAi (double stranded RNA interference) acts at a post-transcriptional level by inducing sequence-specific degradation of homologous mRNAs (Montgomery, Xu and Fire, 1998). Under this aspect, dsRNAi and quelling in *Neurospora* are similar mechanisms, both of them acting at a post-transcriptional level. In addition, both RNA-induced silencing and DNA-induced silencing can be transmitted from cell to cell.

Therefore the identification of *Neurospora* genes which are involved in the silencing is the first step to modulate the same process in plants, animals and fungi. The silencing modulation is of great relevance when  
5 transgenic organisms able to express the desired phenotype are produced.

The authors of the present invention have already isolated *Neurospora crassa* strains mutated at essential functions for gene silencing (Cogoni and Macino, 1997);  
10 15 independent isolated mutants define three complementation groups, thus identifying the *qde-1*, *qde-2* and *qde-3* genes (*qde* stands for "quelling"-deficient), whose products are essential to the silencing machinery. *qde* genes are essential to the *Neurospora* silencing, as  
15 suggested by the fact that silencing of three independent genes (*al-1*, *al-2* and *qa-2*) is impaired by *qde* mutations (Cogoni and Macino, 1997).

The authors of the present invention have already identified *qde-3* gene (PCT WO 00/327885) and *qde-1* gene  
20 (PCT WO 00/50581).

The authors of the invention have identified and cloned now one out of *Neurospora qde* genes, the *qde-2* gene, thus identifying one of required factors for silencing. By considering the similarity between  
25 "quelling" and co-suppression, genes orthologous to the isolated gene are involved in co-suppression and more generally in gene silencing in other organisms, like plants, fungi and animals.

The present invention can be applied with reference  
30 to two general scopes: 1) silencing potentiation as a tool for inactivating more effectively and durably a

desired gene, and 2) silencing suppression to obtain a better expression of the introduced transgenes.

The isolated *qde-2* gene can be introduced alone or with *qde-1* and/or *qde-3* genes in plants, animals or fungi, in order to inactivate the expression of selected genes. The aim is to activate a sequence-specific silencing mechanism both in deficient organisms and in organisms wherein the same is not very efficient. The gene silencing can be induced also by introducing specific double stranded DNA or RNA sequences, homologous to the gene to be inactivated.

As to the silencing potentiation, the over-expression of one or more genes controlling the phenomenon can lead to higher efficiency and/or stability thereof. Therefore the introduction of *qde-2* gene or of homologous genes thereof in organisms can constitute a tool to repress more effectively gene functions. Particularly this approach is specially useful in plants wherein the co-suppression is usually used for the "knock-out" of gene functions. In plants again the gene silencing potentiation can be used to obtain lines resistant to pathogen virus, by introducing transgenes encoding for viral sequences, in order to achieve the expression inhibition of the virus itself (Flavell et al., 1994).

Analogous applications are suitable for animals, wherein some indications suggest that silencing can inhibit the suitable expression of introduced transgenes (Garrick et al., 1998).

On the contrary, there are instances wherein it is desirable not to have or to reduce the gene silencing, i.e. where a transgene is to be over-expressed. It is

known that the co-suppression is strictly correlated both with the presence of an high copy number of the transgene, and with a transgene high expression. This correlation can hamper the production of transgenic organisms which express a transgene at high levels, because more high is the expression and/or the copy number, more probable is to evoke silencing responses. As above mentioned, analogous mechanisms of gene inactivation, dependent on a high copy number, have been disclosed in animals. In these circumstances plant or animal lines, totally or partially ineffective for silencing, constitute an ideal recipient wherein the desired gene can be over-expressed. The invention can be applied within this scope using different approaches:

15           A) Identification and production of mutant lines in genes homologous to *qde-2* gene, in plants, animals and fungi.

          The identification of *Neurospora qde-2* gene, essential for silencing mechanism, can allow the isolation of mutant lines in other organisms, mutated in genes homologous to *qde-2*. For example by means of amplifications using degenerated primers, designed from the most conserved regions of *qde-2* gene, mutant lines in homologous genes can be identified, by analysis of insertion mutant gene banks, already available for many plant species. Both in fungi and animals such mutants can be obtained, following the identification of the homologous gene, by means of "gene disruption" techniques using homologous recombination.

30           B) Reduction of *qde-2* gene expression

          Other strategies for the production of silencing-deficient lines comprise the use of *Neurospora qde-2* gene



or homologous genes thereof. *qde-2* or homologous genes can be introduced into suitable expression vectors to express them in an anti-sense orientation in order to inhibit the expression of resident endogenous genes.

5 Alternatively portions of *qde-2* or of homologous genes can be over-expressed, in order to obtain a negative dominant effect and thus blocking the function of *qde-2* endogenous genes.

The authors of the present invention have cloned  
10 and characterised the *Neurospora crassa qde-2* gene. The sequence analysis of the *qde-2* gene detected a region having a significant homology with the sequence of a *C. elegans* gene, *rde-1*, involved in the dsRNA mediated interference (Tabara et al., 1999).

15 The authors of the invention for the first time have demonstrated that the transgene induced post-transcriptional gene silencing and the dsRNA interference share common genetic mechanisms. This supports the hypothesis that the sequence specific gene silencing  
20 phenomena evolved from an ancestral mechanism aimed to protect the genome against transposons. Furthermore, the results of the authors suggest that dsRNA molecules are involved in the post-transcriptional gene silencing in fungi. dsRNA molecules could be produced directly from  
25 integrated transgenes as a result of the presence of inverted repeats or as an out come of transcription from convergent inverted promoters. Alternatively, single stranded aberrant RNA may be used as a template by an RNA-dependent RNA polymerase (such as QDE-1 protein) able  
30 to produce dsRNAs.

Within the scope of the invention the term homology is intended as similarity, i.e. number of identical

residues + number of conserved residues with respect to the total residues of the considered sequence.

Therefore it is an object of the present invention an isolated nucleic acid molecule encoding for a protein characterized in having a silencing activity and in comprising a domain responsible for dsRNA interference, wherein the domain is at least 25% homologous with the amino acid sequence from aa. 373 to aa. 910 of sequence in fig. 1 (SEQ ID No. 2). Preferably the domain is at least 30% homologous with the amino acid sequence from aa. 373 to aa. 910 of sequence in fig. 1 (SEQ ID No. 2). More preferably the domain is at least 38% homologous with the amino acid sequence from aa. 373 to aa. 910 of sequence in fig. 1 (SEQ ID No. 2). Most preferably the domain comprises the amino acid sequence from aa. 373 to aa. 910 of sequence in fig. 1 (SEQ ID No. 2). According to a particular embodiment the isolated nucleic acid molecule encodes for a protein having the amino acid sequence of fig. 1 (SEQ ID No. 2) or functional portions thereof. Even more preferably the isolated nucleic acid molecule has the sequence of fig. 1 (SEQ ID No. 1) or its complementary sequence.

A further object of the invention is an expression vector comprising, under the control of a promoter which directs the expression in bacteria, the isolated nucleic acid molecule of the invention. Those skilled in the art will appreciate that any plasmid suitable for a correct and effective expression of the protein of the expression in bacteria can be used and it is within the scope of the invention.

A further object of the invention is an expression vector comprising, under the control of a promoter which

directs the expression in plants or in specific plant organs, the isolated nucleic acid molecule of the invention, both in a sense and anti-sense orientation. Those skilled in the art will appreciate that any plasmid  
5 suitable for a correct and effective expression of the protein of the invention in plants or in specific plant organs can be used and it is within the scope of the invention.

A further object of the invention is an expression  
10 vector comprising, under the control of a promoter which directs the expression in fungi, the isolated nucleic acid molecule of the invention, both in a sense and anti-sense orientation. Those skilled in the art will appreciate that any plasmid suitable for a correct and  
15 effective expression of the inventive protein in fungi can be used and it is within the scope of the invention.

A further object of the invention is an expression vector comprising, under the control of a promoter which directs the expression in animals, the isolated nucleic  
20 acid molecule of the invention, both in a sense and anti-sense orientation. Those skilled in the art will appreciate that any plasmid suitable for a correct and effective expression of the protein of the invention in animals can be used and it is within the scope of the  
25 invention.

A further object of the invention is a prokaryotic organism transformed by using the expression vector active in bacteria of the invention.

A further object of the invention is a plant or a  
30 specific plant organ transformed by using the expression vector active in plants of the invention.

A further object of the invention is a plant mutated at the isolated nucleic acid molecule of the invention having a reduced or inhibited silencing activity.

5 A further object of the invention is a fungus transformed with the expression vector of the invention active in fungi.

A further object of the invention is a fungus mutated at the isolated nucleic acid molecule of the invention and having reduced or inhibited silencing activity.

A further object of the invention is a non-human animal transformed with the expression vector of the invention active in animals.

15 A further object of the invention is a non-human animal mutated at the isolated nucleic acid molecule of the invention and having a reduced or inhibited silencing activity.

A further object of the invention refers to a protein characterized in having a silencing activity and in comprising a domain responsible for dsRNA interference, wherein the domain is at least 25% homologous with the amino acid sequence from aa. 373 to aa. 910 in fig. 1 (SEQ ID No. 2). Preferably the domain is at least 30% homologous with the amino acid sequence from aa. 373 to aa. 910 in fig. 1 (SEQ ID No. 2). More preferably the domain is at least 38% homologous with the amino acid sequence from aa. 373 to aa. 910 in fig. 1 (SEQ ID No. 2). Most preferably the domain comprises the amino acid sequence from aa. 373 to aa. 910 in fig. 1 (SEQ ID No. 2). According to a particular embodiment the isolated nucleic acid molecule encodes for a protein

having the amino acid sequence of fig. 1 (SEQ ID No. 2) or functional portions thereof.

It is within the scope of the present invention the use of the isolated nucleic acid molecule of the invention to modulate gene silencing in plants, animals and fungi.

The present invention now will be described by way of non limiting examples with reference to the following figures:

Figure 1: The isolated nucleic acid molecule of the 5.7 Kb fragment containing the *qde-2* gene and flanking sequences (SEQ ID No.1). The amino acid sequence (SEQ ID No. 2) is shown above the nucleotide sequence.

Figure 2: It is schematically represented the pMXY2 plasmid insertion site, in the 80 mutant, used for insertional mutagenesis and consequent polymorphism of the restriction fragments by mean of DNA southern blot of a WT strain and of 80 and 820 mutant strains by using the entire restored flanking region as probe. The 820 mutant has a complete deletion of the *qde-2* gene.

Figure 3: Multiple alignment, at the conserved region, among *qde-2* and other proteins belonging to ago-elF2C family: *A. thaliana ago-1*; rabbit elF2C; *C. elegans rde-1*. Identical amino acids are shown in bold.

## MATERIALS AND METHODS

### *E. coli* strains

*E. coli* strain HB101 ( $F^-$ , *hdsS20*( $rb^-$ ,  $mb^-$ ), *supE44*, *recA13*, *ara14*, *proA2*, *rspL20*( $str^r$ ), *xyl-5*) was used for cloning.

### *Neurospora crassa* strains and growing conditions

*Neurospora crassa* following strains, supplied by Fungal Genetic Stock Center (FGSC, Dpt. Of Microbiology,

University of Kansas Medical Ctr. Kansas City, KA) were used:

- Wild type (FGSC 987);
- *qa-2/aro9* (FGSC 3957A), (FGSC 3958a).

5           The 6XW strain (Cogoni et al., 1996) was obtained upon transformation of the FGSC 3958a strain with pX16 plasmid (Cogoni et al., 1996). This plasmid contains the *qa-2* gene used as selective marker and the *al-1* coding sequence.

10           The mutant strains M7, M20 (*qde-1*); M10, M11 (*qde-2*); M17, M18 (*qde-3*) are described in Cogoni and Macino, 1997.

          The *qde* mutants were obtained by UV mutagenesis. As recipient the transforming strain (6xw) silenced at the  
15           *albino-1* gene was used. *qde* mutants were selected for their ability to recover a wild type unsilenced phenotype and then classified in three different complementation groups. By analyzing the *al-2* gene quelling frequency all of *qde* used mutants are defective for the general  
20           silencing mechanism.

          Complementation assays with not forced heterocaryons were carried out according to Davis and DeSerres, 1970.

#### Plasmids and libraries

25           The plasmid pMXY2, disclosed in Campbell et al. 1994, used for insertional mutagenesis was obtained from Fungal Genetic Stock Center (FGSC, Dpt. Of Microbiology, University of Kansas Medical Ctr. Kansas City, KA). The plasmid contains the *Bml1* gene (allele responsible of the  
30           benilate drug resistance), that was used as selective marker after transformation. The genomic DNA containing

the *qde-2* gene was isolated from a *N. Crassa* gene library in cosmids. (Cabibbo et al., 1991).

#### *N. crassa* transformation

Spheroplasts were prepared according to the Akins and Lambowitz (1985) protocol.

#### Southern Blot Analysis

Chromosomal DNA was prepared as disclosed by Ireland et al., 1993. 5 µg of genomic DNA were digested and blotted as reported in Maniatis et al.

DNA probes were: a) as to the *al-1* gene the probe is represented by a XbaI-ClaI restriction fragment of pX16 (Cogoni et al., 1996); b) as to the *BmI* gene the probe is represented by the 2.6Kb SalI fragment of pMXY2.

#### Northern Blot Analysis

*N. crassa* total RNA was extracted according to the protocol described by Cogoni et al., 1996. The mycelium was grown for two days at 30°C, then powdered in liquid nitrogen before RNA extraction. For Northern analysis 10 µg of RNA were formaldehyde denatured, electrophoresed on a 1% agarose, 7% formaldehyde gel, and blotted over Hybond N (Amersham) membranes. Hybridization was carried out in 50% formamide in the presence of <sup>32</sup>P labeled DNA probe 1.5x10<sup>6</sup> cpm/ml.

#### RESULTS

#### Isolation of silencing mutant by insertional mutagenesis

Previously a *Neurospora* strain (6XW) wherein the *albino-1* resident gene was steadily silenced was used for UV mutagenesis that brought to the isolation of *qde* ("quelling" deficient) mutants in *N. crassa* induced gene silencing (Cogoni and Mancino 1997).

The 6XW strain shows an albino phenotype due to the lack of carotenoid biosynthesis, as results by the

silencing of the albino 1 gene expression (Schmidhauser et al., 1990). A mutation interfering with the silencing machinery is easily detectable by producing a wild type phenotype (bright orange) of the carotenoid biosynthesis.

5 By means of complementation assays it was possible to establish that *qde* mutants belong to three complementation groups, indicating the presence of three genetic loci involved in the *Neurospora* silencing mechanism. In order to isolate the *qde* genes an

10 insertional mutagenesis was carried out with the 6XW strain, previously used for UV mutagenesis. The insertional mutagenesis was carried out by transforming the 6XW strain with a plasmid, taking advantage of the fact that, after the transformation, plasmids are

15 randomly inserted in the *Neurospora crassa* genome. The mutagenesis was carried out transforming the 6XW silenced strain with pMXY2 (see Materials and Methods) which contains the benilate resistance as selective marker. Transformed strains able to grow in the presence of

20 benilate containing medium and showing a wild type phenotype for the carotenoid biosynthesis were selected. Out of 50.000 isolated independent transformed strains, a benilate resistant strain (80) was isolated, which showed the bright orange phenotype expected for a *qde* gene

25 mutation. In order to verify that the silencing release was effectively due to a *qde* gene mutation and not to the loss of *al-1* transgene copies, the genomic DNA of the strain 80 was extracted and digested with *Sma*I and *Hind*III restriction enzymes. After blotting, DNA was

30 hybridized with a probe corresponding to the coding sequence of *al-1*. The *Sma*I site is present only once in the *al-1* transgene containing plasmid and the digestion



by using said enzyme produces a 5.5Kb fragment corresponding to tandem arrayed *al-1* transgenes, while a 3.1Kb fragment is expected from the resident *al-1* locus. The number of *al-1* transgenic copies present in the 80 strain is comparable to that present in the silenced 6XW strain.

The strain 80 is mutated in *qde-2* gene

The strain 80 was assayed in a heterokaryon assay with a wild type strain and with M7, M20 (*qde-1*) M10, M11 (*qde-2*), M17, M18 (*qde-3*) mutants and with a wild strain (Cogoni and Macino, 1997). As shown in Table 1 the *al-1* gene silencing is restored producing an albino phenotype in all of heterokaryons but M10 and M11. This behavior is consistent with the presence of a *qde-2* gene recessive mutation in the strain 80.

Table 1

Reciprocal heterokaryons among the mutant 80 and previously characterized *qde* mutants.

	80	M7	M20	M10	M11	M17	M18
80	WT	AL	AL	WT	WT	AL	AL
M7		WT	WT	AL	AL	AL	AL
M20			WT	AL	AL	AL	AL
M10				WT	WT	AL	AL
M11					WT	AL	AL
M17						WT	WT
M18							WT

WT = heterokaryon with a wild type phenotype for carotenoid accumulation;

AL = heterokaryon with an albino phenotype wherein the *al-1* gene silencing is restored.

Recovery of sequences flanking the pMXY2 plasmid integration site

In order to recover sequences flanking the integration site or sites the following methodology was carried out. The genomic DNA of strain 80 was digested with Aat II enzyme. Subsequently the genomic DNA was ligated and the product used to transform *E. coli* cells that was screened in an ampicillin-containing medium. PQc1 plasmid was recovered and a DNA fragment containing sequences flanking the integration site was isolated from it by using Aat II and Cla I enzymes.

10 Isolation of genomic clones, their subcloning and complementation of the *qde-2* mutant

The fragment from pQc1 plasmid was used to probe a *Neurospora crassa* genomic library in cosmids. Three cosmids 6G10, 20C1 and 23F2 containing about 35 Kb genomic DNA inserts, were isolated. Such cosmids were used in transformation experiments of M11 and 80 mutants. All of cosmids are able to restore the *al-1* gene silencing in the two mutants, determining the appearance of an albino phenotype. The 20C1 cosmid was used to subclone a 5.7 Kb BamHI-BamHI fragment. This subclone was used for transformation experiments and resulted to be able to complement the *qde-2* phenotype, indicating that a *qde-2* functional gene is present in this plasmid.

20 Isolation and sequence of the *qde-2* cDNA

25 The sequence of BamHI-BamHI region allowed to deduce the amino acid sequence of the QDE-2 protein. The *qde-2* gene encodes for a 938 aa. putative protein (104 KDa). The genomic clone does not contain any introns since the reading frame does not contain any interruptions and intron acceptor and donor sequences were not identified (Fig. 1, Seq. ID No 1, 2).

30

The *qde-2* gene comprises an homologous domain with encoding genes for proteins that are responsible for dsRNA interference

5 The 938 aa sequence (SEQ ID No. 2) was used to search in database of amino acid sequences, by using the BLASTP algorithm. As showed in fig. 3, the search identified significant homologies with *argonaute-1* gene [with expected values (E value) of 2e-57] of *A. Thaliana* (mutants of this gene show developmental anomalies); *rde-*  
10 *1* gene [with expected values (E value) of 1e-23] of *C. elegans*, involved in gene silencing phenomena induced by double stranded RNA; *elF2C* gene [with expected values (E value) of 5e-60] of rabbit isolated as an element belonging to transcription beginning complex.

15 Plant expression vector

The *qde-2* gene was inserted, in a sense orientation, into a vector containing a plant expression "cassette", including the 35S promoter and the PI-II "terminator" sequences. The vector also includes the  
20 *Streptomyces hygroscopicus bar* gene, which confers the phosphinotricine herbicide resistance to transformed plants. In an analogous vector to the above mentioned one, *qde-2* was inserted in an anti-sense orientation with respect to the 35S promoter.

25 The obtained vectors can be utilized to over-express the *qde-2* gene in plants, or to repress the gene expression of resident genes, which are homologous to *qde-2*.

Fungus expression vector

30 The *qde-2* gene was inserted in a vector containing a fungal specific expression "cassette", comprising the *A. nidulans trpC* gene promoter and terminator, both in a

sense and an anti-sense orientation. In addition the vector contains the bacterial *hph* gene, which confers the hygromicine drug resistance. The sense plasmid can be used to over express the *qde-2* gene, whereas the anti-sense plasmid is used to repress the expression of *qde-2* homologous genes in various fungine species.

#### Mammalian expression vector

The *qde-2* gene was inserted in a vector containing a mammalian specific expression "cassette", including the cytomegalovirus (CMV) promoter and SV40 termination and polyadenylation sequences both in a sense and anti-sense orientation. The vector includes also the neomycine phototransferase gene, as marker for mammalian cell selection. The sense plasmid can be used to over express the *qde-2* gene, whereas the anti-sense plasmid can be used to repress the expression of *qde-2* homologous genes in various mammalian species.

#### Bibliography

- Akins, R.A. and Lambowitz A.M. (1985) Mol. Cell. Biol. 5:2272-2278
- Baulcombe, D.C. (1996) Plant Mol. Biol. 32, 79-88.
- Cabibbo, A. et al. (1991) Fungal Genetic Newsl., 38: 68-70.
- Campbell J.W. et al. (1994) Fungal Genetic Newsl., 41: 20.
- Cogoni, C. et al. (1996) EMBO J. 15, 3153-3163
- Cogoni, C. and Macino, G. (1997) Proc. Natl. Acad. Sci. U.S.A. 94: 10223-10238.
- Davis, R.H. and De Serres, F.J. (1970) methods Enzymol. 17: 79-143.

- de Carvalho Niebel, F. et al. (1995), Plant Cell : 347-358.
- Dehio, C., and Schell, J. (1994). Proc. Natl. Acad. Sci. U.S.A. 91: 5538-5542.
- 5 - Dorer, D.R. and Henikoff, S. (1994). Cell, 993-1002.
- Elkind, Y. Et al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87: 9057-9061.
- Fire, A. (1999) Trends Genet. 15:358-363.
- Flavell, R.B. (1994) Proc. Natl. Acad. Sci. U.S.A. 91: 3490-3496.
- 10 - Garrick D., et al. (1998) Nature Genetics 18, 56-59.
- Ireland, J. et al. (1993) Fungal Genetics Newsl. 40: 24.
- Maniatis, S.T. et al. (1982) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- 15 - Mittelstein Scheid, O. Et al. (1994) Mol. Gen. Genet. 244: 325-330.
- Montgomery, M.K., Xu, S. and Fire, A. (1998) Proc. Natl. Acad. Sci. USA 95, 15502-7.
- 20 - Pal-Bhadra, M., et al., (1997). Cell 90, 479-490.
- Palauqui, J.C. et al., (1997) EMBO J. 16: 4738-4745.
- Pirrotta, V. (1997). TIG 13, 314-318.
- Schmidhauser, T.J. et al., Mol. Cell. Biol. 10: 5064-5070
- 25 - Stam, M. et al. (1997) Annals of Botany 79:3-12
- Stam, M. et al. (1997) Plant Journal 1:63-82 79:3-12
- Tabara et al. (1999), Cell 99:123-132
- van Blokland, R. et al. (1994), Plant, 6, 861-887.
- 30 - Vaucheret, H. (1993), C.R. Acad. Sci. Paris, Sciences de la vie/Life sciences 316, 1471-1483.

- Wallrath, L.L. and Elgin, S.C.R. (1995). Genes & Development 9, 1263-1277.

## Claims

1. An isolated nucleic acid molecule encoding for a  
5 protein characterized in having a silencing activity and  
in comprising a domain responsible for dsRNA  
interference, wherein the domain is at least 25%  
homologous with the amino acid sequence from aa. 373 to  
aa. 910 of SEQ ID No. 2.

10 2. An isolated nucleic acid molecule encoding for a  
protein characterized in having a silencing activity and  
comprising a domain responsible for dsRNA interference  
according to claim 1, wherein the domain is at least 30%  
homologous with the amino acid sequence from aa. 373 to  
15 aa. 910 of SEQ ID No. 2.

3. An isolated nucleic acid molecule encoding for a  
protein characterized in having a silencing activity and  
comprising a domain responsible for dsRNA interference  
according to claim 2, wherein the domain is at least 38%  
20 homologous with the amino acid sequence from aa. 373 to  
aa. 910 of SEQ ID No. 2.

4. An isolated nucleic acid molecule encoding for a  
protein characterized in having a silencing activity and  
comprising a domain responsible for dsRNA interference  
25 according to claim 3, wherein the domain is the amino  
acid sequence from aa. 373 to aa. 910 of SEQ ID No. 2.

5. An isolated nucleic acid molecule encoding for a  
protein characterized in having a silencing activity and  
comprising a domain responsible for dsRNA interference  
30 according to claim 4, wherein said isolated nucleic acid  
molecule encodes for a protein having the amino acid  
sequence of SEQ ID No. 2, or functional portions thereof.

6. An isolated nucleic acid molecule encoding for a protein characterized in having a silencing activity and comprising a domain responsible for dsRNA interference according to claim 5, wherein said isolated nucleic acid molecule has the sequence of SEQ ID No. 1 or its complementary sequence.

7. Expression vector comprising, under the control of a promoter that directs the expression in bacteria, the isolated nucleic acid molecule according to any one of claims 1-6.

8. Expression vector comprising, under the control of a promoter that directs the expression in plants or in specific plant organs, the isolated nucleic acid molecule according to any one of claims 1-6, both in a sense and anti-sense orientation.

9. Expression vector comprising, under the control of a promoter that directs the expression in fungi, the isolated nucleic acid molecule according to any one of claims 1-6 both in a sense and anti-sense orientation.

10. Expression vector comprising, under the control of a promoter that directs the expression in animals, the isolated nucleic acid molecule according to any one of claims 1-6 both in a sense and anti-sense orientation.

11. Prokaryotic organism transformed by using the expression vector active in bacteria according to claim 7.

12. Plants or a specific plant organ transformed by using the expression vector active in plants according to claim 8.

13. Plant mutated at the isolated nucleic acid molecule according to any one of claims 1-6 having a reduced or inhibited silencing activity.



14. Fungus transformed by using the expression vector active in fungi according to claim 9.

15. Fungus mutated at the isolated nucleic acid molecule according to any one of claims 1-6 having a  
5 reduced or inhibited silencing activity.

16. Non-human animal transformed by using the expression vector active in animals according to claim 10.

17. Non-human animal mutated at the isolated  
10 nucleic acid molecule according to any one of claims 1-6 having a reduced or inhibited silencing activity.

18. Protein characterized in having a silencing activity and comprising a domain responsible for dsRNA interference wherein the domain is at least 25%  
15 homologous to the amino acid sequence from aa. 373 to aa. 910 of SEQ ID No. 2.

19. Protein characterized in having a silencing activity and comprising a domain responsible for dsRNA interference according to claim 18 wherein the domain is  
20 at least 30% homologous to the amino acid sequence from aa. 373 to aa. 910 of SEQ ID No. 2.

20. Protein characterized in having a silencing activity and comprising a domain responsible for dsRNA interference according to claim 19 wherein the domain is  
25 at least 38% homologous to the amino acid sequence from aa. 373 to aa. 910 of SEQ ID No. 2.

21. Protein characterized in having a silencing activity and comprising a domain responsible for dsRNA interference according to claim 20 wherein the domain is  
30 the amino acid sequence from aa. 373 to aa. 910 of SEQ ID No. 2.

22. Protein characterized in having a silencing activity and comprising a domain responsible for dsRNA interference according to claim 21 comprising the amino acid sequence of SEQ ID No. 2 or functional portions thereof.

23. Use of the isolated nucleic acid molecule according to any one of claims 1-6 to modulate the gene silencing in plants, animals and fungi.

1/7

Length of cBAMqde2.txt: 5746 bp; Listed from: 1 to: 5746;  
 Translated from: 1039 to: 3852 (ORFs);  
 Genetic Code used: Universal; Lun, 27 ago 1956 18:50

## Frame 1

```

GGA TCC GCG TAG CAC ATC CTT TTC TTT TCC TTT TGG TTA TCC ATA ACC TTG GCA ACA CCT
      9      18      27      36      45      54

TTC TTT GCT TTC TCT CTC TTT TTC GCT TTA GAG ACC TAC GCA ACT ACC CAT CAT CAT TTT CTG ATA
      69      78      87      96     105     114     123

TCG ACA TAT CAC CCA ACA ACA TCA TCA TCT ACT ACC AGT AAT CCC GCA TCG GAG GAG TAG TCG
     135     144     153     162     171     180     189

TTT CGC TCG ATT ACT CTT TTT TTT GCG TCC GGA GTG CGA CAA AGT AGC GGC TTA TAA CAA GTC CAA
     201     210     219     228     237     246     255

GTT GGA AAA AAA CCA TCA ATC AGT GGT ATT TCT CTC TTG GCA AAT CCA CAA CAA TCC CCT TCC ACG
     267     276     285     294     303     312     321

ACA AAC AAA CAA ACA ACC TAC CTT AAC TAT CCT CTT GCT TAC CTA CGT ACC TGC CTA CCT ACC TAC
     333     342     351     360     369     378     387

CTA CCT ACC TAC CTC TGC TCA ACC AAC CAT CTC GTC AAT CAA ACC GAA CCG AAC CAA ACC GAA CGA
     399     408     417     426     435     444     453

TAG CCG AAT AAG CTC TCG TGC CTT GTT GCT CTA CTC GAC AAT CTG TTA CCA CCA ACA CTA CAA GTT
     465     474     483     492     501     510     519

TAA CAG TCA TGT CTG ACA ATC GTG GCG GTC GTG GAG GTC GTG GCG GCG GTG GTC GCG GCG GCG GCG
     531     540     549     558     567     576     585

GCG GCG GCG GAG GCC GTG GAG GTG GTC AGC AAG GCG GCG GTG GAG GCC GTG GAG GTG GTT ACC AAG
     597     606     615     624     633     642     651

GCA GCG GCG GCG GTG GAG GCC GTG GCG GCG GTT ATC AAG GCG GTG GCG GCG GTG ACC GTG GAG GCC
     663     672     681     690     699     708     717

GTG GCG GCG GTT ATC AAG GCG GTG GTG GCG GTG GTT TCC AAG GCG GCG GTG GAA GGG GTG GCC GTG
     729     738     747     756     765     774     783

GCG GCG GTT TCC AAG GCG GCG GCG GCG GCG GCC GTG GTG GCT TCG GCG GAG GAC AGG GCG CGG GAG
     795     804     813     822     831     840     849

GAT ACG AAC CCC CTC CAC CGG ATG TCT ACA AGT AGG TGC CTC TCC ATT TTT TTT TAC CAT TCA ACA
     861     870     879     888     897     906     915

TGA TGC TGA CAC GAC TTT AGG GGA ATT GAC GGT CGT GGT GCC CCC GAG CCT GAC GCC CAG ATC ACC
     927     936     945     954     963     972     981

                                     M   S   K   L
AAA CTC GAG GAT GAT TGG ATC AAG AAG CAC GTC AGC GAC AAT CTG GTC ACT TCC ATG AGC AAG CTT
     993     1002     1011     1020     1029     1038     1047

S   L   S   E   K   E   K   A   N   N   L   P   V   R   P   G   H   G   T   M   G   E
TCG CTC AGC GAG AAG CAG AAA GCC AAC AAC TTG CCG GTT CGC CCT GGC CAT GGT ACC ATG GGC GAG
     1059     1068     1077     1086     1095     1104     1113

K   V   K   L   W   A   N   Y   F   K   I   N   I   K   S   P   A   I   Y   R   Y   T
AAG GTG AAG CTT TGG GCC AAC TAT TTC AAA ATC AAC ATC AAA TCA CCA GCC ATT TAC AGG TAC ACC
     1125     1134     1143     1152     1161     1170     1179

I   K   V   A   A   T   E   E   K   L   G   K   E   A   E   V   A   S   K   K   V   E
ATC AAA GTT GCC GCC ACC GAG GAA AAG CTC GGA AAG GAA GCT GAG GTC GCA TCC AAG AAA GTG GAG
     1191     1200     1209     1218     1227     1236     1245

V   V   V   G   K   L   L   K   Q   I   E   A   N   V   K   S   V   A   I   A   S   D
GTG GTG GTT GGG AAA CTG CTC AAG CAG ATC GAA GCC AAC GTG AAA TCC GTG GCG ATT GCC AGC GAT
     1257     1266     1275     1284     1293     1302     1311

```

FIG. 11

2/7

F K V H L V T T T K L K V P E N R I F E V T  
 TTC AAA GTG CAC CTG GTG ACG ACC ACC AAG CTC AAA GTT CCC GAG AAC CGC ATC TTT GAG GTG ACG  
 1323 1332 1341 1350 1359 1368 1377  
 W T E P S S N Q N L P S K P Q T W V V K V E  
 TGG ACC GAG CCG AGT TCC AAC CAA AAC CTG CCC AGC AAG CCC CAG ACT TGG GTG GTC AAG GTG GAG  
 1389 1398 1407 1416 1425 1434 1443  
 E S V E T C D F G K V L N E L T T L D P K L  
 GAG AGT GTC GAA ACC TGC GAT TTC GGC AAG GTG CTG AAC GAG CTC ACG ACA CTT GAT CCC AAG CTC  
 1455 1464 1473 1482 1491 1500 1509  
 D G D F P K Y N V E L D A L N T I V T H H A  
 GAC GGA GAC TTT CCC AAG TAC AAT GTG GAG CTC GAT GCC CTC AAC ACC ATT GTG ACT CAT CAT GCC  
 1521 1530 1539 1548 1557 1566 1575  
 R A D D N V A V V G R G R F F A I G D D L I  
 CGC GCC GAC GAC AAT GTT GCG GTG GTG GGA AGG GGA AGG TTT TTT GCC ATT GGT GAT GAC CTC ATT  
 1587 1596 1605 1614 1623 1632 1641  
 E Q V R P H D S P L V I L R G Y F A S V R P  
 GAA CAA GTG CGG CCC CAT GAC TCC CCT TTG GTC ATC TTG CGA GGA TAT TTT GCC AGC GTC CGT CCA  
 1653 1662 1671 1680 1689 1698 1707  
 A T G R L L L N T N I T H G V F R P G V K L  
 GCT ACC GGA AGA CTT TTA CTC AAT ACC AAC ATC ACG CAT GGT GTC TTC CGT CCT GGG GTC AAA CTT  
 1719 1728 1737 1746 1755 1764 1773  
 A Q L F Q E L G L D V M D K C N A W N E V T  
 GCA CAG CTG TTT CAG GAA CTT GGA CTT GAC GTA ATG GAC AAA TGC AAT GCC TGG AAC GAA GTA ACC  
 1785 1794 1803 1812 1821 1830 1839  
 K N Q L N D K M R R V H K V L A K G R V E L  
 AAA AAT CAG CTC AAC GAC AAG ATG CGC AGA GTT CAC AAG GTC CTG GCT AAG GGC CGT GTC GAG TTG  
 1851 1860 1869 1878 1887 1896 1905  
 N A P F L I D G K I V Y K K C Y R T L N G I  
 AAT GCC CCA TTC CTT ATT GAT GGA AAG ATT GTT TAT AAA AAA TGT TAC CGC ACG CTC AAT GGC ATT  
 1917 1926 1935 1944 1953 1962 1971  
 A N R G D E R G K Q K D G K E V R Y P P L F  
 GCT AAC CGT GGC GAC GAA AGG GGG AAG CAA AAG GAT GGT AAA GAA GTC CGA TAT CCG CCC TTG TTC  
 1983 1992 2001 2010 2019 2028 2037  
 G I P G V Q V G G P T S C Q F Y L R A R E T  
 GGG ATT CCG GGT GTC CAG GTT GGC GGC CCG ACC TCT TGT CAG TTC TAC TTG CGT GCG CGA GAG ACA  
 2049 2058 2067 2076 2085 2094 2103  
 K D G A A P P P T P G L P S N A Y I T V A N  
 AAG GAT GGC GCT GCC CCT CCT CCG ACT CCC GGC CTG CCG AGC AAC GCG TAC ATC ACG GTA GCG AAC  
 2115 2124 2133 2142 2151 2160 2169  
 Y Y K Q R Y G I T A N A S L P L V N V G T K  
 TAT TAT AAA CAA CGG TAC GGA ATA ACC GCC AAT GCT TCG CTT CCT CTG GTC AAC GTT GGC ACC AAG  
 2181 2190 2199 2208 2217 2226 2235  
 E K A I Y V L A E F C T L V K G R S V K A K  
 GAA AAG GCG ATT TAC GTC TTG GCC GAG TTT TGT ACG CTG GTC AAA GGC CGT TCC GTC AAG GCT AAG  
 2247 2256 2265 2274 2283 2292 2301  
 L T A N E A D N M I K F A C R A P S L N A Q  
 CTG ACA GCC AAC GAG GCG GAC AAC ATG ATT AAG TTT GCT TGC AGA GCT CCT TCG CTG AAC GCT CAG  
 2313 2322 2331 2340 2349 2358 2367  
 S I V T K G R Q T L G L D K S L T L G K F K  
 TCT ATC GTG ACG AAA GGC AGA CAG ACA CTT GGT CTT GAT AAA AGC CTG ACG CTT GGC AAG TTC AAG  
 2379 2388 2397 2406 2415 2424 2433  
 V S I D K E L I T V V G R E L K P P M L T Y  
 GTT TCG ATC GAC AAG GAG CTG ATC ACC GTT GTC GGG CGT GAG CTC AAG CCT CCG ATG CTT ACG TAC  
 2445 2454 2463 2472 2481 2490 2499

FIG. 1-2

3/7

S G N K T V E P Q D G G W L M K F V K V A R  
 AGC GGT AAC AAG ACG GTA GAG CCG CAG GAC GGC GGG TGG TTG ATG AAG TTT GTC AAG GTC GCC AGA  
 2511 2520 2529 2538 2547 2556 2565  
 P C R K I E K W T Y L E L K G S K A N E G V  
 CCT TGC CGC AAG ATT GAG AAG TGG ACA TAC TTG GAA CTG AAG GGT TCC AAG GCA AAC GAA GGG GTG  
 2577 2586 2595 2604 2613 2622 2631  
 P Q A M T A F A E F L N R T G I P I N P R F  
 CCG CAA GCT ATG ACC GCT TTT GCC GAA TTC TTG AAC AGA ACG GGC ATC CCG ATT AAC CCC AGG TTC  
 2643 2652 2661 2670 2679 2688 2697  
 S P G M S M S V P G S E K E F F A K V K E L  
 TCG CCG GGC ATG AGC ATG TCA GTT CCA GGG AGC GAA AAA GAG TTC TTT GCC AAA GTG AAG GAA CTC  
 2709 2718 2727 2736 2745 2754 2763  
 M S S H Q F V V V L L P R K D V A I Y N M V  
 ATG AGC TCG CAC CAA TTT GTG GTG GTT CTT TTA CCC AGA AAG GAT GTT GCG ATC TAC AAT ATG GTG  
 2775 2784 2793 2802 2811 2820 2829  
 K R A A D I T F G V H T V C C V A E K F L S  
 AAG CGG GCT GCC GAT ATC ACA TTT GGC GTT CAC ACA GTC TGT TGT GTA GCC GAA AAG TTC CTT AGC  
 2841 2850 2859 2868 2877 2886 2895  
 T K G Q L G Y F A N V G L K V N L K F G G T  
 ACT AAG GGG CAG CTG GGG TAT TTT GCC AAC GTC GGC CTC AAG GTC AAC CTC AAG TTT GGC GGC ACC  
 2907 2916 2925 2934 2943 2952 2961  
 N H N I K T P I P L L A K G K T M V V G Y D  
 AAT CAC AAT ATC AAG ACG CCC ATT CCT TTG CTC GCC AAG GGG AAG ACG ATG GTG GTG GGC TAT GAT  
 2973 2982 2991 3000 3009 3018 3027  
 V T H P T N L A A G Q S P A S A P S I V G L  
 GTC ACC CAT CCG ACC AAT CTA GCG GCT GGA CAA TCG CCT GCA TCG GCT CCC AGT ATT GTC GGC CTG  
 3039 3048 3057 3066 3075 3084 3093  
 V S T I D Q H L G Q W P A M V W N N P H G Q  
 GTC TCA CAT ATC GAC CAA CAC CTT GGA CAA TGG CCT GCA ATG GTT TGG AAC AAC CCG CAC GGC CAG  
 3105 3114 3123 3132 3141 3150 3159  
 E S M T E Q F T D K F K T R L E L W R S N P  
 GAG TCC ATG ACG GAA CAG TTT ACG GAC AAG TTC AAG ACG CGT CTG GAA CTA TGG CGC AGC AAT CCC  
 3171 3180 3189 3198 3207 3216 3225  
 A N N R S L P E N I L I F R D G V S E G Q F  
 GCA AAC AAC CGC AGT CTC CCC GAG AAT ATC CTG ATT TTC CGC GAT GGC GTC TCC GAG GGA CAG TTC  
 3237 3246 3255 3264 3273 3282 3291  
 Q M V I K D E L P L V R A A C K L V Y P A G  
 CAG ATG GTC ATC AAG GAC GAG CTA CCC CTG GTT CGC GCC GCC TGC AAG CTG GTG TAT CCA GCT GGC  
 3303 3312 3321 3330 3339 3348 3357  
 K L P R I T L I V S V K R H Q T R F F P T D  
 AAG CTA CCG CGT ATT ACG CTG ATT GTC TCT GTC AAG CGC CAC CAG ACT CGC TTC TTC CCA ACG GAC  
 3369 3378 3387 3396 3405 3414 3423  
 P K H I H F K S K S P K E G T V V D R G V T  
 CCG AAG CAT ATT CAC TCC AAG TCC AAG AGC CCC AAG GAG GGT ACT GTG GTT GAC CGC GGC GTG ACC  
 3435 3444 3453 3462 3471 3480 3489  
 N V R Y W D F F L Q A H A S L Q G T A R S A  
 AAC GTC CGC TAT TGG GAC TTC TTT TTG CAG GCG CAC GCG TCG CTC CAG GGC ACG GCC CGC TCG GCT  
 3501 3510 3519 3528 3537 3546 3555  
 H Y T V L V D E I F R A D Y G N K A A D T L  
 CAC TAC ACA GTT CTG GTG GAT GAG ATT TTC AGG GCC GAC TAT GGA AAC AAG GCG GCC GAC ACG CTG  
 3567 3576 3585 3594 3603 3612 3621  
 E Q L T H D M C Y L F G R A T K A V S I C P  
 GAG CAG CTG ACG CAT GAC ATG TGT TAT CTC TTT GGA CGA GCC ACC AAG GCT GTC AGT ATC TGC CCG

FIG. 1-3

4/7

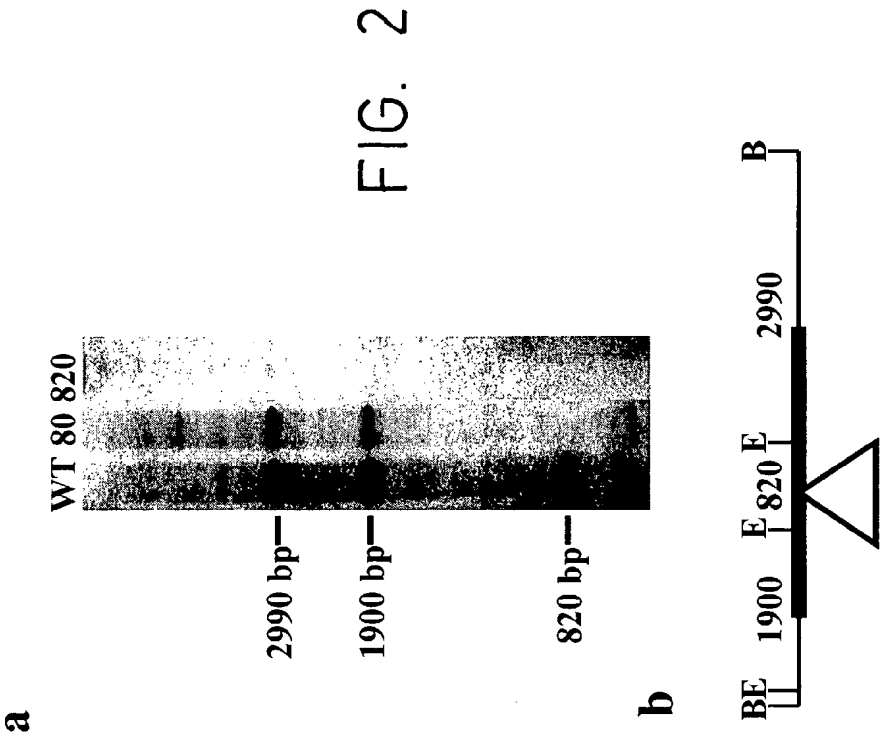
3633	3642	3651	3660	3669	3678	3687
P A Y Y A D L V C D R A R I H Q K E L F D A						
CCT GCG TAC TAT GCC GAC TTG GTG TGC GAC CGG GCG CGT ATC CAT CAG AAG GAG CTC TTT GAC GCC						
3699	3708	3717	3726	3735	3744	3753
L D E N D S V K T D D F A R W G N S G A V H						
CTC GAT GAA AAC GAT AGC GTT AAG ACC GAT GAT TTC GCA AGA TGG GGT AAC TCC GGG GCT GTT CAT						
3765	3774	3783	3792	3801	3810	3819
P N L R N S M Y Y I						
CCC AAC CTT AGG AAC TCC ATG TAC TAT ATC TAG GCT TGT CAA TTG TGT GCT GGA ATG TAC TGG AGC						
3831	3840	3849	3858	3867	3876	3885
ATA TAA GTG ACG CGA TGG AAG CCT AAT CGT CTC TGA ATA TGG ATC AAA GAC AGC GTT TGC TTT TTC						
3897	3906	3915	3924	3933	3942	3951
GGG GCT TCT AGT TTC TAC AGC GAT TTG TGT GGA TTG TTT CTT GTT CTG TTT CTT GGT TCT TTC TTT						
3963	3972	3981	3990	3999	4008	4017
CTT TTT TTT GTG TCT CTG TCT GCC TTT GTA CTG CAT GCA AAC GTG CAC TCT GAA TGA TGA ACG ACA						
4029	4038	4047	4056	4065	4074	4083
CCA TTT GAC GAT TGG ATA AGA GAT GAC AGA CTG CAG ATA CTA TCA TGC GCA ATG GAA AAC ACG AAC						
4095	4104	4113	4122	4131	4140	4149
AAC CAA GGT TTT TGA TTC CTT CAA TAG CGA AAT ATA GAA AAA GAA ACA AAA AAA AAA ACA ACA ACA						
4161	4170	4179	4188	4197	4206	4215
AAT AAT GGA AGT ATG ATT AAA CAC ATT GAG CGC GAT GAC TGA CTG GTG TTG TGA ATG GCG TGT TGG						
4227	4236	4245	4254	4263	4272	4281
TTT TCT TCT TTC TTG AAA ATT TAG AAC CGT AAA TGT TAT ATC ATG TGA TGT AAT GTA ATA ACA TAT						
4293	4302	4311	4320	4329	4338	4347
TTA TAT CTC GTT GTA TTC TTG TAC ACA CTT TCC AGG ATA ACA TGG TCT GAC ATG GTA TTT CTG ACG						
4359	4368	4377	4386	4395	4404	4413
TAC AAA AAA GAA AAA GAA AAA CAG GAA ACC ATG AAC CCG CGA CAA AGC TGT TCC AGT TGT TAC AAT						
4425	4434	4443	4452	4461	4470	4479
GAT GAT GAT GAT GAT GAC CTA CTA CCT AAG GTA TTC TAT CTT AGC CAA GGT ATT CTC TCG CAT CCT						
4491	4500	4509	4518	4527	4536	4545
ATT CCA TCC TAT CCT AAC CCG AGC CTA ACC CGA GCC TAA ATA CCT AAA CTC CTA AAC TCC TTA ACT						
4557	4566	4575	4584	4593	4602	4611
CCT TAA CTC CTT TCT AAA TGT CTA AAC CCC CAA ACT ATG AGA CGA CCC GAA CCC GAA ACC CTA ATA						
4623	4632	4641	4650	4659	4668	4677
AAA GTA TTT ATA AAC CAT CAT AAA AGA AAA AAA ACC ATC ATA CAT GGA TGA TCA AAA CAA ACA GAA						
4689	4698	4707	4716	4725	4734	4743
ACG GAA ACA ACA CAA CCA GCT ACC CGC TCA AGA CTT TCA TTC GTT AAT TCA TCA CTC ACT CAC TCA						
4755	4764	4773	4782	4791	4800	4809
CTC ACT CAC TCA GCA GCA AAA TAC CGT TTT GTC CTG CTA TTC GTT TGT TGC GCC TTG ATT TCA GGC						
4821	4830	4839	4848	4857	4866	4875
GGG ACA ATG GTG TGA TGT ACG ACG TGG GGG CGG TAG ACT GCG TCT ACT GGT GGC ATC CTT TAC AAT						
4887	4896	4905	4914	4923	4932	4941
TTT TTA GTG TGT CAG TAT GTG ATG TAT TCA ATG CTA TTG AAC TGA GGG GGG CTG ATG GAT AGT GGG						
4953	4962	4971	4980	4989	4998	5007
GAG AGA ACA CCT GAC GGA TAG AGG GAA GGA ACT GGA CGC CTG GGG GGA AGT GAG AGA GGG GGA TGG						
5019	5028	5037	5046	5055	5064	5073
TGG GGA ATA GAT GAA AAG AGA AGA GGA GTG AGA GCA CAA GAA GAA AGA ATG AAT GTT GGT GAC AAA						
5085	5094	5103	5112	5121	5130	5139

FIG. 1-4

5/7

GTT AAA GAA AAG GAA GGG GGG AAA GAG AAG AGG ACA GGT GTG GTG AGT GAA TTG AGT GAA AGG AAG  
 5151 5160 5169 5178 5187 5196 5205  
 GGA AAA AAC GGA GAA GGA AAA AAA AAA CAT AAA AAA AAA AAA AAA AAC AGA AAG AAA GAA CTA ACC  
 5217 5226 5235 5244 5253 5262 5271  
 AAT CAT CCA AAC TCA GCG GAA AGT ACT CAT ACA AAA GGT CGG CTG CCT CAA TCG GAC TCC CCA CAT  
 5283 5292 5301 5310 5319 5328 5337  
 TCT CTT TCT GGT ACT GAT TCT GCT GCC CCA GAC TTC CAC TTT CAA AGT GGC TAT CAC CCT TAT TGT  
 5349 5358 5367 5376 5385 5394 5403  
 TGT TAG AGT GAG TAG TAG ACG TAA GTC CTC CCG ATC CGG AGC CAA AAC CCA TCC CTT TCC CAG CTG  
 5415 5424 5433 5442 5451 5460 5469  
 TAT CCC TCT TCA ATC CAC CAG TAG CAA CAC CCA TCT TGC CAT AGA GCG GAC TAT CCC CTG CCC CTG  
 5481 5490 5499 5508 5517 5526 5535  
 CCC CTG CCG AGC CAG GAG TAG CAG TCC TAT TCA TAG GCG GAC TCC TCT GCT CGT CTT CCG ACA GGG  
 5547 5556 5565 5574 5583 5592 5601  
 ACA AAC TAA TTG GTA GGG CAC CCG CAG CAG AGG AGG AGG TAT TTC TGT GAT GAC TGG TTC TGT TTG  
 5613 5622 5631 5640 5649 5658 5667  
 GGG CAG CTA AGG GCG TGG GTT TCC TTC GTG AGC CGC TGT TGT GAT TGT TGG CGG CGG CGT CCG AGG  
 5679 5688 5697 5706 5715 5724 5733  
 ATA AGG ATC C  
 5745

FIG. 1-5





7/7

AGO-1 LGSRLPAYDGRKSLTAGPLPFNSKEFRINLLDEEYAGCQRREREFKVVKLVARADLH 300  
 eIF2C FGDRKPVFDGRKSLTAMPLPIGRE----KVELEVTLPEGKDRIFKVSFKWVSCVSLQ 95  
 QDE-2 LTLDEKLDQDFPKNVVELDALNTIVTHHARADDNVAVVCRGRFFAIGDDLTIEQVRPHDS 210  
 RDE-1 ILTYRKKFHLNFSRENPEKDEEANR---SYKFLKNVMTQKVRYAPFVNEEIRVQFAKNFV 216

AGO-1 HIGMFLEGKQSDAPQALQVQDIVRELPTSRYIPVGRSEYSPDIGKKQSLGDLGLESWRG 360  
 eIF2C ALHDALSGRLPSVPFETIQADVVVRHLPSMRYTPVGRSEFTASEGCSNPLGGGRVWFG 155  
 QDE-2 PLVILRGYFASVRBATGRLLNTNTHGVFRPGVKLAQLQELGLDVMCKNAWNEVTKN 270  
 RDE-1 YDNNSILRVESFHDPNRFEQSLEVPAPRIEAWFGIYIGIKELFDGEPVLNFAIVDKLFYN 276

AGO-1 FYQSIRPTOMGLSLN--IDMSSTAFIEANPVIQFVCDLNRDISS--RPLSDADRVKIK 415  
 eIF2C FHQSVRPSLWKMMLN--IDVSATAFYKAQPVIEFVCEVDFKSLIEEQKPLTDSQVVFET 213  
 QDE-2 QLNDKMRVHKVLAKGRVELNAPFLIDGKIYKKCYRTNGIANRGDERGKQKDGKGVRY 330  
 RDE-1 APKMSLLDYLLIVD--PQSCNDDVRKDLKTKLMAGKMTIRQAARPRIQLENLKLCA 334

AGO-1 KALRGVKVEVHRGNMRRKVRISGLTAVATRELTFVDEERN---TQKSVVEYFHETYGFR 472  
 eIF2C KEIKGLKVEIHCGQMRRKRVNCVRRPASHQTFELQOESGQVECTVACQFKDRHKLK 273  
 QDE-2 PPLFGIPGVQVGGPTSCQFYLAARETKDGAAPPPTGLPSN---AYITVANYKYQRYGIT 387  
 RDE-1 EVWDNEMSRLEERHLTFDLDCEENSLVYKVTKGSDRGRNAK---KYDTLTKIYENKKEF 391

AGO-1 IQHTOLECLOVCNSNPNVLEMEVCKVEECQRYSKRLNERQITALLKVTQOREID-REK 530  
 eIF2C LRYPHLPCLQVCQEQKHTYLELVQCNIVAC-QRCIKKLTNDQSTMTIRATARSAPP-RQE 331  
 QDE-2 AN-ASLEPLVNVCTKEKATVLAFFCTVAKC-RSVKAKITANEADNMKFACRASLSNAQS 445  
 RDE-1 IEFPHLEPLVVKSGAKAYAVPMHELVHEKQRYKNRIDLVMQDKFLKRAIRKPHDYKEN 451

AGO-1 DILQTVQLNDYAKDN-YAQEFCIKISTSLASVEARILPPEWLKWHESGREGTCLPOVEOM 589  
 eIF2C EISKLMSASFNTDP-YVREFCIMVKDEMTDVGRVLOPESILNGGRNK-AIATEVQCVN 389  
 QDE-2 IYTKGRQTLGLDKSL-TLGKSKVSIDKELIITVGRBLKPEMLTMSGNKT--VE-RODCGW 501  
 RDE-1 TIRMLKELDFSSEELNFVERFCICSKLOMIECFCKVLEKEMVNSVNEQIKMTPVIRFQ 511

AGO-1 NMMNKKMIN-GGTVNNWICINF--RQVQDNLARTSCQBLAQMCYVSCMAFNPEPVLPPV 646  
 eIF2C DMRNKQFHT-GIEIKVWALACFAPQRQCTEVHLKSETEQIRKISRDAQMPHQQQPCFCYK 448  
 QDE-2 LMKFVKVARPCRKIEKWTYLEK--GSKANEGVPQAMTAFAEFLNRTGIPINPRFSPGMS 559  
 RDE-1 EKQLNVVPEKELCCAEVUVNETAGNPCLEENDVVKYTBETIGGCKFRGIRIGANENRGAQ 571

AGO-1 SARPEQVEKVLKTRY-----HDATSKISQGEIDLLIIVLPDNNGSL 688  
 eIF2C AQGADSVGPMPFR-----HUKNTYAGLQLVVVLPGKTPV 482  
 QDE-2 MSVEGSEKEFFAK-----VKELMSSHQFVVVLPRKDVAI 594  
 RDE-1 SIMYDATKNBYAFYKNCTLNTGIGRFEIAATEAKNMFERIPDKEQKVLMPFIISKRQLNA 631

AGO-1 YGDLKRICETELGVSQCCLTKVFMKSKOY-----MANVALKINVKVGGENTVLV 739  
 eIF2C YAEVVRVGDVTLGATQCVQMNQVORTTPT-----LSNCLKINVKLGGVNNILL 533  
 QDE-2 YNMVKRAADITGCVHTVCCVAKFLSTKGLGY-----FANVGLKINVKFGGNTNHEK 647  
 RDE-1 YGFVKHYCDHTIGVANQHITSETVTKALASLRHEKSGKRIFYQALKINAKLGCINQELD 691

AGO-1 DALSRRIPIFCAOVTHPH-----PGEDSSPSIAAVVASQDWPEITK 788  
 eIF2C PQGR--P-----PVFQQPVIFLCADVTHPP-----AGDGKIPSIAAVVCMDAHPN-R 578  
 QDE-2 TPIPI-----LLAKGKTIVVGVDTVHTNLAAGQSPASPSIVGLVSTIDQHLG-Q 696  
 RDE-1 WSEIAEISPEEKERRKTMPLTMYVCHDVTHPT-----SYSGIDYSIAAVVASINPGGT-I 745

AGO-1 YAGLVCAQAHROELIQDLFEKWKDPQKGVVTGGMIKELLIAFRSTGH-KPLRIEFYRDG 847  
 eIF2C YCATVRVQCHROELIQDLAAMVR-----ELLIQFYKSTRE-KPTRLIEFYRDG 624  
 QDE-2 WPAMVWNNPHCOESMTEQFTDKFKTR-----LELWRSNPANNRS-LEPNTLIEFYRDG 746  
 RDE-1 YRNMIVTQEECRPGERAVAHGRERTD---ILEAKFVRLREFAEENNDRAPAHIVYVRDG 802

AGO-1 VSEGQFYQVLLYELDAIRKACASLEAG-----YQPPVTFVIVQKRHHTRLEAQNHNDRHS 902  
 eIF2C VSEGQFYQVLLHHELDAIREACIKLEK-----YQPGITEFIVQKRHHTRLECTDKNER-- 677  
 QDE-2 VSEGQFYQVLIKDELPLVRAACKLVYPAG-----KLPRITLIVSVKRHHQTRFPTDPKHIH- 801  
 RDE-1 VSDSEMLRVSHDELRSLKSEVKQFMSERDGEDPEPKYTFIVIQKRHHTRLELRMEKDKFPV 862

AGO-1 V-----DRSCNILECTVVDKSKTCHPTEFDFYLCSHAG 934  
 eIF2C V-----GKSGNIPAGTTVDTKITHTEFDFYLCSHAG 709  
 QDE-2 -----FKSKSPKEGTVDVDRGVTVNVRYDFDLQAHAS 832  
 RDE-1 VNKDLTPAETDVAVAAVKQWEEDMKESKETGIVNPSSGTTVDKLVSKYKDFDFLASHHC 922

AGO-1 IQGTSREAHYHVLWDENN-----FTADGLQSLTNLCTYIARCTRSVSIIVPAYIAHLA 988  
 eIF2C IQGTSREASHYHVLWDENNR-----FSSDELQILTYQLCHTYVRCRTSVSIPAPAYIAHLV 763  
 QDE-2 IQGTARSAHYHVLVIEIFRADYGNKAADLEQLTHDMCYLFGRAIKRAVSIICPPAYIADIV 892  
 RDE-1 VLGTSRGCHYTYMYDDKG-----MSQDEVYKTYGIAFLSARCRKPSLFEVVEVYAHLS 976

AGO-1 AFRARFYMEPETSDSGSMASGSMARGGGMAGRSTRGPNVNAAVRPLPALKENVKRVMSYC 1048  
 eIF2C AFRARYHLVDKEHDS-----AEGSHTSQSNGRDHQALAKAVQVQDHLRTLVYFA- 813  
 QDE-2 CDRARIHQKELFDALD-----END-SVKTDDEFARWGNSCAVHPNLRNSMYVI- 938  
 RDE-1 CERAKELYRITYKEHYIG-----DYAQPRTRHEMEHFLQTNVKYPOMSFA- 1020

FIG. 3

## SEQUENCE LISTING

<110> Università degli Studi di Roma La Sapienza

Cogoni, Carlo

Macino, Giuseppe

Catalanotto, Caterina

Azzalin, Gianluca

<120> Isolation and characterization of a N. crassa silencing  
gene and uses thereof

<130> qde-2

<140>

<141>

<150> RM2000A000021

<151> 2000-01-17

<160> 2

<170> PatentIn Ver. 2.1

<210> 1

<211> 5746

<212> DNA

<213> Neurospora crassa

<220>

<221> CDS

<222> (1039)..(3852)

<400> 1

ggatcgcggt agcacatcct tttcttttcc ttttggttat ccataacott ggcaacacct 60  
ttctttgctt tctctctctt ttctgcttta gagacctacg caactaccca tcatcatctt 120  
ctgatatcga catatcaccc aacaacatca tcatcatcta ctaccagtaa tcccgcacgc 180  
gaggagtagt cgtttcgctc gattactctt ttttttgcgt ccggagtgcg acaaagtagc 240  
ggcttataac aagtccaagt tggaaaaaaa ccatcaatca gtggtatttc tctcttgcca 300  
aatccacaac aatccccttc cagacaaaac aaacaaacaa cctaccttaa ctatcctctt 360  
gcttacctac gtacctgcct acctacctac ctacctacct acctctgctc aaccaaccat 420  
ctcgtcaatc aaaccgaacc gaaccaaacc gaacgatagc cgaataagct ctcgtgcctt 480

2/12

Asn Arg Ile Phe Glu Val Thr Trp Thr Glu Pro Ser Ser Asn Gln Asn	
110 115 120	
ctg ccc agc aag ccc cag act tgg gtg gtc aag gtg gag gag agt gtc	1455
Leu Pro Ser Lys Pro Gln Thr Trp Val Val Lys Val Glu Glu Ser Val	
125 130 135	
gaa acc tgc gat ttc ggc aag gtg ctg aac gag ctc acg aca ctt gat	1503
Glu Thr Cys Asp Phe Gly Lys Val Leu Asn Glu Leu Thr Thr Leu Asp	
140 145 150 155	
ccc aag ctc gac gga gac ttt ccc aag tac aat gtg gag ctc gat gcc	1551
Pro Lys Leu Asp Gly Asp Phe Pro Lys Tyr Asn Val Glu Leu Asp Ala	
160 165 170	
ctc aac acc att gtg act cat cat gcc cgc gcc gac gac aat gtt gcg	1599
Leu Asn Thr Ile Val Thr His His Ala Arg Ala Asp Asp Asn Val Ala	
175 180 185	
gtg gtg gga agg gga agg ttt ttt gcc att ggt gat gac ctc att gaa	1647
Val Val Gly Arg Gly Arg Phe Phe Ala Ile Gly Asp Asp Leu Ile Glu	
190 195 200	
caa gtg cgg ccc cat gac tcc cct ttg gtc atc ttg cga gga tat ttt	1695
Gln Val Arg Pro His Asp Ser Pro Leu Val Ile Leu Arg Gly Tyr Phe	
205 210 215	
gcc agc gtc cgt cca gct acc gga aga ctt tta ctc aat acc aac atc	1743
Ala Ser Val Arg Pro Ala Thr Gly Arg Leu Leu Leu Asn Thr Asn Ile	
220 225 230 235	
acg cat ggt gtc ttc cgt cct ggg gtc aaa ctt gca cag ctg ttt cag	1791
Thr His Gly Val Phe Arg Pro Gly Val Lys Leu Ala Gln Leu Phe Gln	
240 245 250	
gaa ctt gga ctt gac gta atg gac aaa tgc aat gcc tgg aac gaa gta	1839
Glu Leu Gly Leu Asp Val Met Asp Lys Cys Asn Ala Trp Asn Glu Val	
255 260 265	
acc aaa aat cag ctc aac gac aag atg cgc aga gtt cac aag gtc ctg	1887
Thr Lys Asn Gln Leu Asn Asp Lys Met Arg Arg Val His Lys Val Leu	
270 275 280	
gct aag ggc cgt gtc gag ttg aat gcc cca ttc ctt att gat gga aag	1935
Ala Lys Gly Arg Val Glu Leu Asn Ala Pro Phe Leu Ile Asp Gly Lys	
285 290 295	
att gtt tat aaa aaa tgt tac cgc acg ctc aat ggc att gct aac cgt	1983

Ile Val Tyr Lys Lys Cys Tyr Arg Thr Leu Asn Gly Ile Ala Asn Arg	
300 305 310 315	
ggc gac gaa agg ggg aag caa aag gat ggt aaa gaa gtc cga tat ccg	2031
Gly Asp Glu Arg Gly Lys Gln Lys Asp Gly Lys Glu Val Arg Tyr Pro	
320 325 330	
ccc ttg ttc ggg att ccg ggt gtc cag gtt ggc ggc ccg acc tct tgt	2079
Pro Leu Phe Gly Ile Pro Gly Val Gln Val Gly Gly Pro Thr Ser Cys	
335 340 345	
cag ttc tac ttg cgt gcg cga gag aca aag gat ggc gct gcc cct cct	2127
Gln Phe Tyr Leu Arg Ala Arg Glu Thr Lys Asp Gly Ala Ala Pro Pro	
350 355 360	
ccg act ccc ggc ctg ccg agc aac gcg tac atc acg gta gcg aac tat	2175
Pro Thr Pro Gly Leu Pro Ser Asn Ala Tyr Ile Thr Val Ala Asn Tyr	
365 370 375	
tat aaa caa cgg tac gga ata acc gcc aat gct tcg ctt cct ctg gtc	2223
Tyr Lys Gln Arg Tyr Gly Ile Thr Ala Asn Ala Ser Leu Pro Leu Val	
380 385 390 395	
aac gtt ggc acc aag gaa aag gcg att tac gtc ttg gcc gag ttt tgt	2271
Asn Val Gly Thr Lys Glu Lys Ala Ile Tyr Val Leu Ala Glu Phe Cys	
400 405 410	
acg ctg gtc aaa ggc cgt tcc gtc aag gct aag ctg aca gcc aac gag	2319
Thr Leu Val Lys Gly Arg Ser Val Lys Ala Lys Leu Thr Ala Asn Glu	
415 420 425	
gcg gac aac atg att aag ttt gct tgc aga gct cct tcg ctg aac gct	2367
Ala Asp Asn Met Ile Lys Phe Ala Cys Arg Ala Pro Ser Leu Asn Ala	
430 435 440	
cag tct atc gtg acg aaa ggc aga cag aca ctt ggt ctt gat aaa agc	2415
Gln Ser Ile Val Thr Lys Gly Arg Gln Thr Leu Gly Leu Asp Lys Ser	
445 450 455	
ctg acg ctt ggc aag ttc aag gtt tcg atc gac aag gag ctg atc acc	2463
Leu Thr Leu Gly Lys Phe Lys Val Ser Ile Asp Lys Glu Leu Ile Thr	
460 465 470 475	
gtt gtc ggg cgt gag ctc aag cct ccg atg ctt acg tac agc ggt aac	2511
Val Val Gly Arg Glu Leu Lys Pro Pro Met Leu Thr Tyr Ser Gly Asn	
480 485 490	
aag acg gta gag ccg cag gac ggc ggg tgg ttg atg aag ttt gtc aag	2559

Lys Thr Val Glu Pro Gln Asp Gly Gly Trp Leu Met Lys Phe Val Lys	
495 500 505	
gtc gcc aga cct tgc cgc aag att gag aag tgg aca tac ttg gaa ctg	2607
Val Ala Arg Pro Cys Arg Lys Ile Glu Lys Trp Thr Tyr Leu Glu Leu	
510 515 520	
aag ggt tcc aag gca aac gaa ggg gtg ccg caa gct atg acc gct ttt	2655
Lys Gly Ser Lys Ala Asn Glu Gly Val Pro Gln Ala Met Thr Ala Phe	
525 530 535	
gcc gaa ttc ttg aac aga acg ggc atc ccg att aac ccc agg ttc tcg	2703
Ala Glu Phe Leu Asn Arg Thr Gly Ile Pro Ile Asn Pro Arg Phe Ser	
540 545 550 555	
ccg ggc atg agc atg tca gtt cca ggg agc gaa aaa gag ttc ttt gcc	2751
Pro Gly Met Ser Met Ser Val Pro Gly Ser Glu Lys Glu Phe Phe Ala	
560 565 570	
aaa gtg aag gaa ctc atg agc tcg cac caa ttt gtg gtg gtt ctt tta	2799
Lys Val Lys Glu Leu Met Ser Ser His Gln Phe Val Val Val Leu Leu	
575 580 585	
ccc aga aag gat gtt gcg atc tac aat atg gtg aag cgg gct gcc gat	2847
Pro Arg Lys Asp Val Ala Ile Tyr Asn Met Val Lys Arg Ala Ala Asp	
590 595 600	
atc aca ttt ggc gtt cac aca gtc tgt tgt gta gcc gaa aag ttc ctt	2895
Ile Thr Phe Gly Val His Thr Val Cys Cys Val Ala Glu Lys Phe Leu	
605 610 615	
agc act aag ggg cag ctg ggg tat ttt gcc aac gtc ggc ctc aag gtc	2943
Ser Thr Lys Gly Gln Leu Gly Tyr Phe Ala Asn Val Gly Leu Lys Val	
620 625 630 635	
aac ctc aag ttt ggc ggc acc aat cac aat atc aag acg ccc att cct	2991
Asn Leu Lys Phe Gly Gly Thr Asn His Asn Ile Lys Thr Pro Ile Pro	
640 645 650	
ttg ctc gcc aag ggg aag acg atg gtg gtg ggc tat gat gtc acc cat	3039
Leu Leu Ala Lys Gly Lys Thr Met Val Val Gly Tyr Asp Val Thr His	
655 660 665	
ccg acc aat cta gcg gct gga caa tcg cct gca tcg gct ccc agt att	3087
Pro Thr Asn Leu Ala Ala Gly Gln Ser Pro Ala Ser Ala Pro Ser Ile	
670 675 680	
gtc ggc ctg gtc tca acc atc gac caa cac ctt gga caa tgg cct gca	3135

Val Gly Leu Val Ser Thr Ile Asp Gln His Leu Gly Gln Trp Pro Ala			
685	690	695	
atg gtt tgg aac aac ccg cac ggc cag gag tcc atg acg gaa cag ttt	3183		
Met Val Trp Asn Asn Pro His Gly Gln Glu Ser Met Thr Glu Gln Phe			
700	705	710	715
acg gac aag ttc aag acg cgt ctg gaa cta tgg cgc agc aat ccc gca	3231		
Thr Asp Lys Phe Lys Thr Arg Leu Glu Leu Trp Arg Ser Asn Pro Ala			
720	725	730	
aac aac cgc agt ctc ccc gag aat atc ctg att ttc cgc gat ggc gtc	3279		
Asn Asn Arg Ser Leu Pro Glu Asn Ile Leu Ile Phe Arg Asp Gly Val			
735	740	745	
tcc gag gga cag ttc cag atg gtc atc aag gac gag cta ccc ctg gtt	3327		
Ser Glu Gly Gln Phe Gln Met Val Ile Lys Asp Glu Leu Pro Leu Val			
750	755	760	
cgc gcc gcc tgc aag ctg gtg tat cca gct ggc aag cta ccg cgt att	3375		
Arg Ala Ala Cys Lys Leu Val Tyr Pro Ala Gly Lys Leu Pro Arg Ile			
765	770	775	
acg ctg att gtc tct gtc aag cgc cac cag act cgc ttc ttc cca acg	3423		
Thr Leu Ile Val Ser Val Lys Arg His Gln Thr Arg Phe Phe Pro Thr			
780	785	790	795
gac ccg aag cat att cac ttc aag tcc aag agc ccc aag gag ggt act	3471		
Asp Pro Lys His Ile His Phe Lys Ser Lys Ser Pro Lys Glu Gly Thr			
800	805	810	
gtg gtt gac cgc ggc gtg acc aac gtc cgc tat tgg gac ttc ttt ttg	3519		
Val Val Asp Arg Gly Val Thr Asn Val Arg Tyr Trp Asp Phe Phe Leu			
815	820	825	
cag gcg cac gcg tcg ctc cag ggc acg gcc cgc tcg gct cac tac aca	3567		
Gln Ala His Ala Ser Leu Gln Gly Thr Ala Arg Ser Ala His Tyr Thr			
830	835	840	
gtt ctg gtg gat gag att ttc agg gcc gac tat gga aac aag gcg gcc	3615		
Val Leu Val Asp Glu Ile Phe Arg Ala Asp Tyr Gly Asn Lys Ala Ala			
845	850	855	
gac acg ctg gag cag ctg acg cat gac atg tgt tat ctc ttt gga cga	3663		
Asp Thr Leu Glu Gln Leu Thr His Asp Met Cys Tyr Leu Phe Gly Arg			
860	865	870	875
gcc acc aag gct gtc agt atc tgc ccg cct gcg tac tat gcc gac ttg	3711		

Ala Thr Lys Ala Val Ser Ile Cys Pro Pro Ala Tyr Tyr Ala Asp Leu  
880 885 890

gtg tgc gac cgg gcg cgt atc cat cag aag gag ctc ttt gac gcc ctc 3759  
Val Cys Asp Arg Ala Arg Ile His Gln Lys Glu Leu Phe Asp Ala Leu  
895 900 905

gat gaa aac gat agc gtt aag acc gat gat ttc gca aga tgg ggt aac 3807  
Asp Glu Asn Asp Ser Val Lys Thr Asp Asp Phe Ala Arg Trp Gly Asn  
910 915 920

tcc ggg gct gtt cat ccc aac ctt agg aac tcc atg tac tat atc 3852  
Ser Gly Ala Val His Pro Asn Leu Arg Asn Ser Met Tyr Tyr Ile  
925 930 935

taggcttgctc aattgtgtgc tggaatgtac tggagcatat aagtgacgcg atggaagcct 3912

aatcgtctct gaatatggat caaagacagc gtttgctttt tcggggcttc tagtttctac 3972

agcgatttgt gtggattgtt tcttgttctg tttcttggtt ctttctttct tttttttgtg 4032

tctctgtctg cctttgtact gcatgcaaac gtgcactctg aatgatgaac gacaccattt 4092

gacgattgga taagagatga cagactgcag atactatcat gcgcaatgga aaacacgaac 4152

aaccaagggtt tttgattcct tcaatagcga aatatagaaa aagaaacaaa aaaaaaaca 4212

acaacaaata atggaagtat gattaaacac attgagcgcg atgactgact ggtgttgtga 4272

atggcgtggt ggttttcttc tttcttgaaa atttagaacc gtaaattgta tatcatgtga 4332

tgtaatgtaa taacatatatt atatctcgtt gtattcttgt acacactttc caggataaca 4392

tggtctgaca tggatattct gacgtacaaa aaagaaaaag aaaaacagga aaccatgaac 4452

ccgcgacaaa gctgttcacg ttgttacaat gatgatgatg atgatgacct actacctaag 4512

gtattctatc ttagccaagg tattctctcg catcctattc catcctatcc taacccgagc 4572

ctaaccgag cctaaatacc taaactccta aactccttaa ctccttaact cctttctaaa 4632

tgtctaaacc cccaaactat gagacgaccc gaaccgaaa ccctaataaa agtatttata 4692

aaccatcata aaagaaaaaa aaccatcata catggatgat caaaacaaac agaaacggaa 4752

acaacacaac cagctaccg ctcaagactt tcattcggtta attcatcact cactcactca 4812

ctcactcact cagcagcaaa ataccgtttt gtctgtctat tcgtttgttg cgccttgatt 4872



tcaggcggga caatggtgtg atgtacgacg tgggggcggt agactgcgtc tactggtggc 4932  
 atcctttaca attttttagt gtgtcagtat gtgatgtatt caatgctatt gaactgaggg 4992  
 gggctgatgg atagtgggga gagaacacct gacggataga gggaaggaac tggacgcctg 5052  
 gggggaagtg agagaggggg atggtgggga atagatgaaa agagaagagg agtgagagca 5112  
 caagaagaaa gaatgaatgt tggtgacaaa gttaaagaaa aggaaggggg gaaagagaag 5172  
 aggacaggtg tggtgagtga attgagtga aggaagggaa aaaacggaga aggaaaaaaa 5232  
 aaacataaaa aaaaaaaaaa aaacagaaag aaagaactaa ccaatcatcc aaactcagcg 5292  
 gaaagtactc atacaaaagg tcggctgcct caatcggact cccacattc tctttctggt 5352  
 actgattctg ctgccccaga cttccacttt caaagtggct atcaccctta ttgttgtag 5412  
 agtgagtagt agacgtaagt cctcccgatc cggagccaaa acccatccct tcccagctg 5472  
 tatccctctt caatccacca gtagcaacac ccattctgcc atagagcgga ctatccctg 5532  
 cccctgcccc tgccgagcca ggagtagcag tcctattcat aggcggactc ctctgctcgt 5592  
 cttccgacag ggacaaacta attggtaggg caccgcagc agaggaggag gtatttctgt 5652  
 gatgactggt tctgtttggg gcagctaagg gcgtgggttt ccttcgtgag ccgctgttgt 5712  
 gattgttggc ggcggcgtcc gaggataagg atcc 5746

&lt;210&gt; 2

&lt;211&gt; 938

&lt;212&gt; PRT

&lt;213&gt; Neurospora crassa

&lt;400&gt; 2

Met Ser Lys Leu Ser Leu Ser Glu Lys Glu Lys Ala Asn Asn Leu Pro  
 1 5 10 15

Val Arg Pro Gly His Gly Thr Met Gly Glu Lys Val Lys Leu Trp Ala  
 20 25 30

Asn Tyr Phe Lys Ile Asn Ile Lys Ser Pro Ala Ile Tyr Arg Tyr Thr  
 35 40 45

Ile Lys Val Ala Ala Thr Glu Glu Lys Leu Gly Lys Glu Ala Glu Val

50                      55                      60

Ala Ser Lys Lys Val Glu Val Val Val Gly Lys Leu Leu Lys Gln Ile  
65                      70                      75                      80

Glu Ala Asn Val Lys Ser Val Ala Ile Ala Ser Asp Phe Lys Val His  
                    85                      90                      95

Leu Val Thr Thr Thr Lys Leu Lys Val Pro Glu Asn Arg Ile Phe Glu  
                    100                      105                      110

Val Thr Trp Thr Glu Pro Ser Ser Asn Gln Asn Leu Pro Ser Lys Pro  
                    115                      120                      125

Gln Thr Trp Val Val Lys Val Glu Glu Ser Val Glu Thr Cys Asp Phe  
                    130                      135                      140

Gly Lys Val Leu Asn Glu Leu Thr Thr Leu Asp Pro Lys Leu Asp Gly  
145                      150                      155                      160

Asp Phe Pro Lys Tyr Asn Val Glu Leu Asp Ala Leu Asn Thr Ile Val  
                    165                      170                      175

Thr His His Ala Arg Ala Asp Asp Asn Val Ala Val Val Gly Arg Gly  
                    180                      185                      190

Arg Phe Phe Ala Ile Gly Asp Asp Leu Ile Glu Gln Val Arg Pro His  
                    195                      200                      205

Asp Ser Pro Leu Val Ile Leu Arg Gly Tyr Phe Ala Ser Val Arg Pro  
                    210                      215                      220

Ala Thr Gly Arg Leu Leu Leu Asn Thr Asn Ile Thr His Gly Val Phe  
225                      230                      235                      240

Arg Pro Gly Val Lys Leu Ala Gln Leu Phe Gln Glu Leu Gly Leu Asp  
                    245                      250                      255

Val Met Asp Lys Cys Asn Ala Trp Asn Glu Val Thr Lys Asn Gln Leu  
                    260                      265                      270

Asn Asp Lys Met Arg Arg Val His Lys Val Leu Ala Lys Gly Arg Val  
                    275                      280                      285

Glu Leu Asn Ala Pro Phe Leu Ile Asp Gly Lys Ile Val Tyr Lys Lys  
                    290                      295                      300

Cys Tyr Arg Thr Leu Asn Gly Ile Ala Asn Arg Gly Asp Glu Arg Gly

305		310		315		320
Lys Gln Lys Asp Gly Lys Glu Val Arg Tyr Pro Pro Leu Phe Gly Ile						
	325		330		335	
Pro Gly Val Gln Val Gly Gly Pro Thr Ser Cys Gln Phe Tyr Leu Arg						
	340		345		350	
Ala Arg Glu Thr Lys Asp Gly Ala Ala Pro Pro Pro Thr Pro Gly Leu						
	355		360		365	
Pro Ser Asn Ala Tyr Ile Thr Val Ala Asn Tyr Tyr Lys Gln Arg Tyr						
	370		375		380	
Gly Ile Thr Ala Asn Ala Ser Leu Pro Leu Val Asn Val Gly Thr Lys						
	385		390		395	400
Glu Lys Ala Ile Tyr Val Leu Ala Glu Phe Cys Thr Leu Val Lys Gly						
	405		410		415	
Arg Ser Val Lys Ala Lys Leu Thr Ala Asn Glu Ala Asp Asn Met Ile						
	420		425		430	
Lys Phe Ala Cys Arg Ala Pro Ser Leu Asn Ala Gln Ser Ile Val Thr						
	435		440		445	
Lys Gly Arg Gln Thr Leu Gly Leu Asp Lys Ser Leu Thr Leu Gly Lys						
	450		455		460	
Phe Lys Val Ser Ile Asp Lys Glu Leu Ile Thr Val Val Gly Arg Glu						
	465		470		475	480
Leu Lys Pro Pro Met Leu Thr Tyr Ser Gly Asn Lys Thr Val Glu Pro						
	485		490		495	
Gln Asp Gly Gly Trp Leu Met Lys Phe Val Lys Val Ala Arg Pro Cys						
	500		505		510	
Arg Lys Ile Glu Lys Trp Thr Tyr Leu Glu Leu Lys Gly Ser Lys Ala						
	515		520		525	
Asn Glu Gly Val Pro Gln Ala Met Thr Ala Phe Ala Glu Phe Leu Asn						
	530		535		540	
Arg Thr Gly Ile Pro Ile Asn Pro Arg Phe Ser Pro Gly Met Ser Met						
	545		550		555	560
Ser Val Pro Gly Ser Glu Lys Glu Phe Phe Ala Lys Val Lys Glu Leu						

565	570	575
Met Ser Ser His Gln Phe Val Val Val Leu Leu Pro Arg Lys Asp Val		
580	585	590
Ala Ile Tyr Asn Met Val Lys Arg Ala Ala Asp Ile Thr Phe Gly Val		
595	600	605
His Thr Val Cys Cys Val Ala Glu Lys Phe Leu Ser Thr Lys Gly Gln		
610	615	620
Leu Gly Tyr Phe Ala Asn Val Gly Leu Lys Val Asn Leu Lys Phe Gly		
625	630	635
Gly Thr Asn His Asn Ile Lys Thr Pro Ile Pro Leu Leu Ala Lys Gly		
645	650	655
Lys Thr Met Val Val Gly Tyr Asp Val Thr His Pro Thr Asn Leu Ala		
660	665	670
Ala Gly Gln Ser Pro Ala Ser Ala Pro Ser Ile Val Gly Leu Val Ser		
675	680	685
Thr Ile Asp Gln His Leu Gly Gln Trp Pro Ala Met Val Trp Asn Asn		
690	695	700
Pro His Gly Gln Glu Ser Met Thr Glu Gln Phe Thr Asp Lys Phe Lys		
705	710	715
Thr Arg Leu Glu Leu Trp Arg Ser Asn Pro Ala Asn Asn Arg Ser Leu		
725	730	735
Pro Glu Asn Ile Leu Ile Phe Arg Asp Gly Val Ser Glu Gly Gln Phe		
740	745	750
Gln Met Val Ile Lys Asp Glu Leu Pro Leu Val Arg Ala Ala Cys Lys		
755	760	765
Leu Val Tyr Pro Ala Gly Lys Leu Pro Arg Ile Thr Leu Ile Val Ser		
770	775	780
Val Lys Arg His Gln Thr Arg Phe Phe Pro Thr Asp Pro Lys His Ile		
785	790	795
His Phe Lys Ser Lys Ser Pro Lys Glu Gly Thr Val Val Asp Arg Gly		
805	810	815
Val Thr Asn Val Arg Tyr Trp Asp Phe Phe Leu Gln Ala His Ala Ser		

820	825	830
Leu Gln Gly Thr Ala Arg Ser Ala His Tyr Thr Val Leu Val Asp Glu		
835	840	845
Ile Phe Arg Ala Asp Tyr Gly Asn Lys Ala Ala Asp Thr Leu Glu Gln		
850	855	860
Leu Thr His Asp Met Cys Tyr Leu Phe Gly Arg Ala Thr Lys Ala Val		
865	870	875 880
Ser Ile Cys Pro Pro Ala Tyr Tyr Ala Asp Leu Val Cys Asp Arg Ala		
	885	890 895
Arg Ile His Gln Lys Glu Leu Phe Asp Ala Leu Asp Glu Asn Asp Ser		
	900	905 910
Val Lys Thr Asp Asp Phe Ala Arg Trp Gly Asn Ser Gly Ala Val His		
	915	920 925
Pro Asn Leu Arg Asn Ser Met Tyr Tyr Ile		
930	935	